

COMPUTER MODELING OF MARINE WATERS WITH PUBLIC DOMAIN SOFTWARE

Bert Rubash and Elizabeth Kilanowski
Raincoast GeoResearch
P.O. Box 2206, Bellingham, WA 98227
kilaruba@copper.net

Keywords: open-source, circulation, modeling, COHERENS, contaminant, Bellingham Bay

Abstract

Scientific computing developed historically within international scientific traditions governed by copyright law. By contrast, commercial property rights and patent law have regulated other forms of computing—engineering, financial, entertainment, publishing, and personal computing. Within the scientific tradition, collaboration and citation are sufficient motivation and reward, and national and international cooperation is a catalyst and sometimes a necessity. Within the commercial tradition, exclusionary rights, royalties and license fees are necessary rewards, and co-workers in national and international communities are competitors.

Programming tools, the Linux operating system, computational, visualization and analysis software distributed under the General Public License (GPL) and similar public domain licenses have created an opportunity to develop inexpensive modeling projects that produce sophisticated results using software resources and methods developed within the scientific computing tradition. Two important advantages to using public domain software are financial independence that allows neglected projects to be chosen, and the openness that permits extensive, voluntary collaboration.

The use of efficient and effective, freely available software is illustrated by a project to model circulation and tracer transport in Bellingham Bay in northern Puget Sound. The use of a sampling of public domain software is outlined and illustrated. Its adaptability, maintainability and superior lifespan is discussed.

1 Objectives

1.1 Independence and objectivity

Independent research, independent of funding constraints and funding controls, is able to choose its object and approach it with an objectivity that is beyond the reach of research that must define itself as a project that will attract funding. Whenever economic forces have nothing to gain or something to lose from the results of a research project, a large reservoir of funding will be unavailable for that project. Moreover, many of the remaining sources of funds are subject to influence or direct pressure from those who would rather see money spent on research from which they might gain, or who simply do not want money spent on research that may have results contrary to their interests.

Among those whose interests may cause them to prefer estuarine research that is less than objective are developers, home buyers, aquaculturists, loggers, farmers, public utilities, paper mills, and military organizations.

1.2 Contribute to and benefit from scientific traditions and resources

1.2.1 The structure of modeling

The development of a model can be analyzed into four parts, or four steps in a cycle: when a modeling cycle is finished it provides the basis for another, more mature cycle.

Physical principles and laws The starting place for any type of modeling—computer modeling, mathematical modeling, or simply conceptual modeling—is the collection of perceptual abstractions that we call science that humanity has accumulated over the last few millennia. Scientific abstraction is the means by which we organize and understand our experience of the physical world, and is often the basis of perception itself.

Mathematical representations and methods Translating abstract principles and laws into usable mathematical expressions is a process of evaluating compromises and simplifications, and making judgments, assumptions, and interpretations. This step is the beginning of the practical modeling process.

Numerical methods One of the assumptions of most hydrodynamic theory is smooth continuity, that is, that the atoms and molecules that make up a fluid are so small that continuous mathematics is the most effective means to perform the computations of fluid dynamics. Unfortunately, computers can only perform the computations of discrete mathematics. A field of study called numerical methods that is dedicated to solving this difficulty appeared long before computers arrived, but is now a rapidly developing science. Application of the principles and methods of numerical methods, like the steps above, also requires understanding compromises and making judgments.

Computer science, programming methods A final step includes the tasks of choosing computer programming resources and methods that satisfy the computer resource-intensive needs of fluid dynamics computations, of uncovering the unavoidable errors in extensive and complicated source code, of configuring a model for particular physical conditions and uses, and of writing code for analysis and verification.

The cycle The intellectual cycle begins with doubt which drives the transition from physical reality to conceptual abstraction. The cycle then returns to experience by means of mathematical abstraction and practical application. Practical application is where the cycle begins anew with a better grasp of physical reality, and sometimes with new perceptions.

1.2.2 Common ownership and collaboration versus property rights and exclusion

At the dawn of the computer age and for sometime after, following the traditions of open scientific collaboration, all software source code was shared and in the public domain. Sometime around the second decade, and with increasing extent and force, property rights invaded the cycle presented above, starting at the last step. The consequences of the arrival of property rights may be less destructive once hydrodynamic modeling has achieved maturity and becomes a normal engineering profession. Such modeling, however, although a powerful tool in the hands of scientists, is still a relatively undeveloped scientific method and much collaborative work remains to be done.

By monopolizing both the ownership of a model and information about the details of the methods used, intellectual property rights block the essential cycle of science except within the monopolizing organization, and only inasmuch as the cycle promotes the interests of the organization.

1.2.3 Traditional tools

UNIX is the operating system running the computer workstations on which most modeling research and development has been done. Much of the scientific software in the public domain that has been developed in the past and is currently being developed on workstations has not been adapted to run on other operating systems. For that reason a UNIX operating system provides the widest choice and availability of software for modeling projects.

1.2.4 Reliability and adaptability

Software in the public domain, including development tools and operating systems, can, in effect, be owned by users. Users can fix problems, adapt and optimize software to their uses and needs; they can share their fixes and adaptations, and use the fixes and adaptations of their colleagues. Users of private domain software are invariably restricted in their uses, and depend on market forces to allocate resources to make bug fixes and adaptations.

2 An illustrative example: Modeling Bellingham Bay

2.1 Operating system

Two public versions of the UNIX system, FreeBSD and GNU/Linux, are available for inexpensive personal computers, PC's. The Gentoo Linux distribution was chosen because the distribution and upgrades could be downloaded without cost from the Internet, but more importantly because it would be compiled with optimizations set for the particular hardware it would control. A system optimized for particular computer hardware is a feature available only with public operating systems for PC's. Commercial, closed source systems are necessarily one-size-fits-all systems.

Linux is generally purchased or downloaded from the Web as a *distribution*, which is a collection of software including the GNU/Linux kernel, the X-windows windowing system, several choices of window manager, and hundreds of support and utility programs.

2.2 Modeling program

The Princeton Ocean Model, POM, [Blumberg & Mellor, 1987] developed by Alan Blumberg and George Mellor in the late 1970's and continuously improved since through worldwide collaboration with modelers and users, is the most widely used and emulated ocean circulation model. POM is licensed under the GPL.

POM's development cycle is an embodiment of the open development cycle outlined above. The cycle is manifest as a robust collection of modeling projects and scientific journal articles, and an active online discussion. The open and public nature of the cycle is guaranteed by the GPL licensing of the source code.

COHERENS, modeling software that employs much that was learned in the development of POM, was chosen for the Bellingham Bay project. COHERENS was developed over the period 1990-1998 by a multinational European group funded by the European Union. [Huthnance, 1997, Luyten, 1999, Proctor, 1997] COHERENS is comprehensive and adaptable, providing the user with many choices of what to model and how to model.

NetCDF capabilities were written into COHERENS for this project. NetCDF (network Common Data Form) is a machine independent format for representing scientific data, created to support the creation, access, and sharing of scientific data. [NetCDF] NetCDF libraries are licensed under the University Corporation for Atmospheric Research/Unidata public license.

2.3 Programming environment

2.3.1 Compiler and debugger

COHERENS source code was written in the Fortran 77 language. All of the Fortran code was translated into C++ for this project in order to improve program structure, to add dynamic memory allocation and other pointer programming capabilities, to replace arrays of variables with array classes using the Blitz++ high performance library, and to add other features of a modern computer language. The translation into C++ also enabled the use of many free programming utilities that make writing and maintaining program code more efficient, effective, and reliable. Blitz++ is licensed under the GPL or, if desired, under the "Blitz++ Artistic License"[Blitz++]. Translation from Fortran was initiated using the public domain program f2c from AT&T Bell Laboratories.

The GNU compiler, GCC is being used in this project for all of program code writing, configuration, and maintenance. It is licensed under the GPL.

GDB, a companion of the GNU compiler, is used to trace program execution and analyze variables. It is also licensed under the GPL.

2.3.2 Text editor

NEdit is used for all text editing. It can edit very large data files, hundreds of megabytes, with efficiency and dexterity. NEdit was originally created at Fermilab and is now maintained and distributed by NEdit.org and licensed under the GPL.[NEdit]

2.3.3 Integrating user interface

Finally, Kdevelop serves to integrate the developer's tools within a graphic interface. It is licensed under the GPL.[Kdevelop]

2.4 Configuration

2.4.1 Bathymetry

The first practical step in model configuration was to design a computational grid indexed to the bathymetry of the region being modeled. Raw soundings data was obtained from NOAA. Figure 1 shows the locations of the 93,562

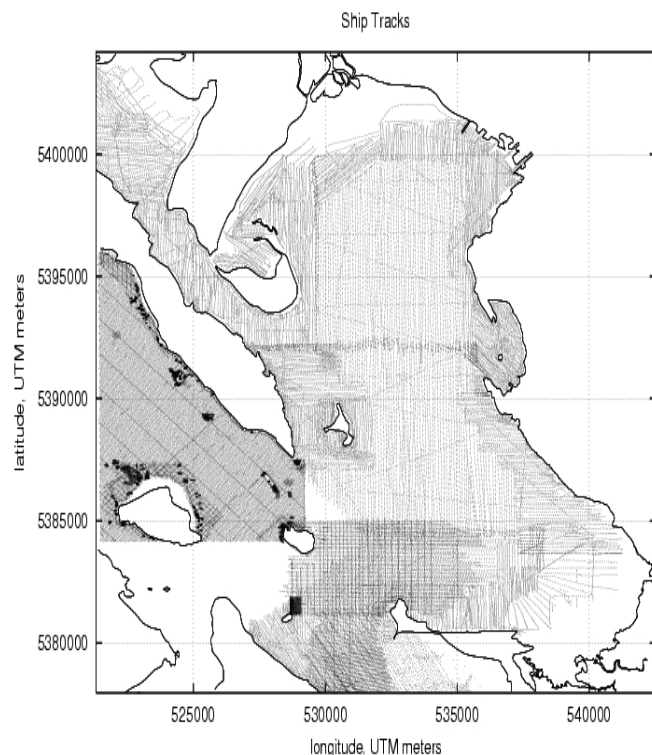


Figure 1: Map of 93,562 soundings from National Ocean Service Hydrographic Survey Data [NOS].

individual soundings used for model configuration. Figure 1 was constructed using Gnuplot[Gnuplot], a general purpose plotting program licensed under the GPL.

The raw bathymetry data was processed by creating a Delauney triangulated surface from the data, and then interpolating from the triangulated surface to a rectangular grid. Programs from Generic Mapping Tools (GMT) performed the processing and plotting (triangulate, grdcontour, grdimage).

Figure 2 is a color-filled contour plot of the triangulated and gridded bathymetry cropped at the open sea boundaries chosen for the model. Cropping was done with the spreadsheet-like xgridedit program from GMT which reads data in the NetCDF format.

2.4.2 Tides

Tide constituents for Bellingham Bay were taken from XTide data files. A single diurnal constituent was used for a tide elevation at the open sea boundaries for test runs of the model. Writing code to couple real-time XTide elevation predictions and tide current predictions at the open boundaries is planned for future work.[XTide]

2.5 Runtime analysis

Several test runs were necessary to verify that boundary conditions determining tide elevations, stream inflows rates, and wind stresses were properly configured. To do this, values of variables of interest were sent to files at selected intervals and then analyzed with Gnuplot. Gnuplot was also used to monitor what is called the Courant–Friedrichs–Levy (CFL) condition, a measure of stability; if time-step increments are too large in relation to fluid velocities and cell dimensions, the CFL condition is violated and modeling fails to produce coherent, or even finite results.

The CFL condition is monitored whenever changes to configuration or modeling methods are made.

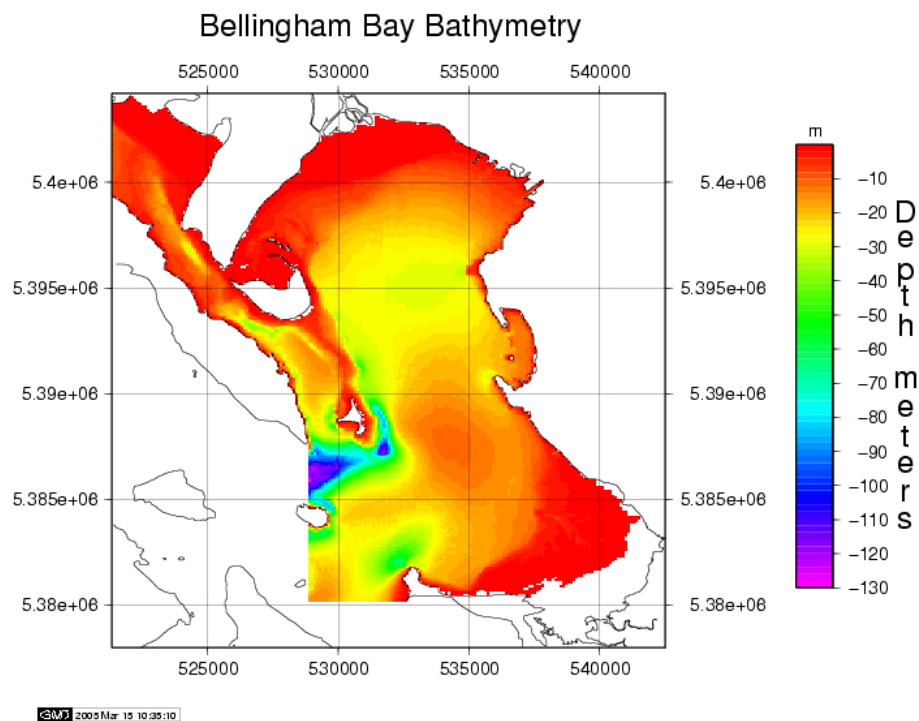


Figure 2: Filled color contour image of gridded bathymetry data.

2.6 Visualization and processing of results

The main visualization products that are used to analyze output are line and color filled contour plots, vector fields, surfaces and volumes. Animations of all these products are made to show time related phenomena such as oscillating patterns, surges, or moving density difference interfaces.

Python is used as a scripting language to edit, reformat, or otherwise process model input, output and boundary conditions.

The visualization and analysis software used includes Gnuplot and GMT, which were already mentioned, Ghostscript, Ncview, and Vis5d+[Gnuplot, GMT, GS, Ncview, Vis5d+].

Figures 3 and 4 are single frames from animations of color coded vector fields and of color filled contours of tracer concentration.

2.7 Report and article preparation

The \LaTeX document typesetting system is used instead of a word processor. \LaTeX documents have long been standard in scientific and academic publishing, and many journals and institutions maintain and distribute \LaTeX templates for contributors. The \LaTeX system and a large collection of accessory packages and templates is freely available software, licensed under the \LaTeX Project Public License, the LPPL. \LaTeX standards have been developed to guarantee that arcane, but essential mathematical expressions and fine, but necessary graphic detail in author submissions appear in final publications.

2.8 Status of the project

Initial configuration and testing is complete. Accurate representation of tides, streamflow, and other conditions are being compiled, and adapted to the model or programmed into the model. Fieldwork for calibration and verification will soon begin. Results of the test runs have inspired additions to the project and new projects.

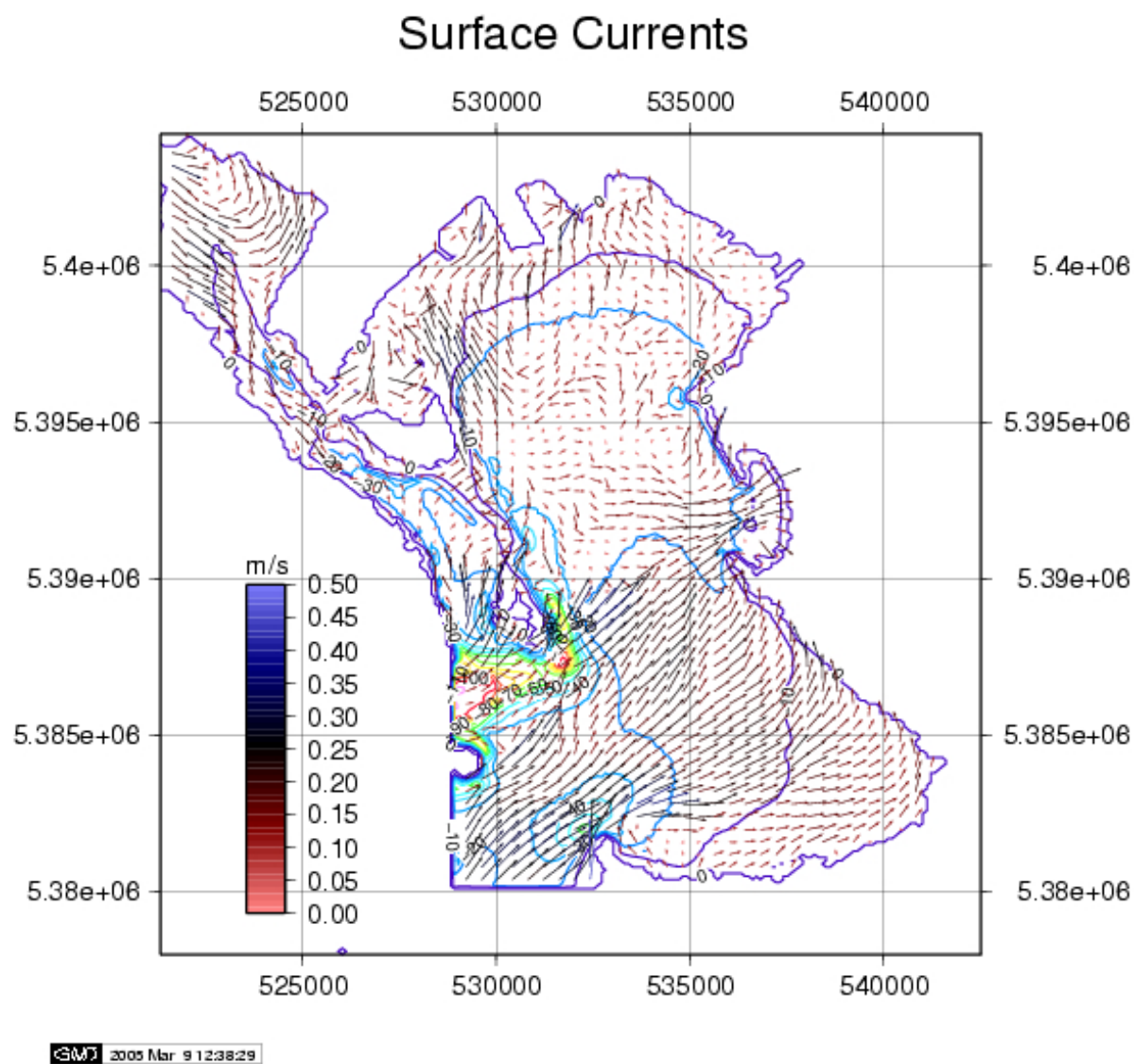


Figure 3: Color coded bathymetry contours and a vector field of surface currents.

Whatcom Creek Tracer Plume

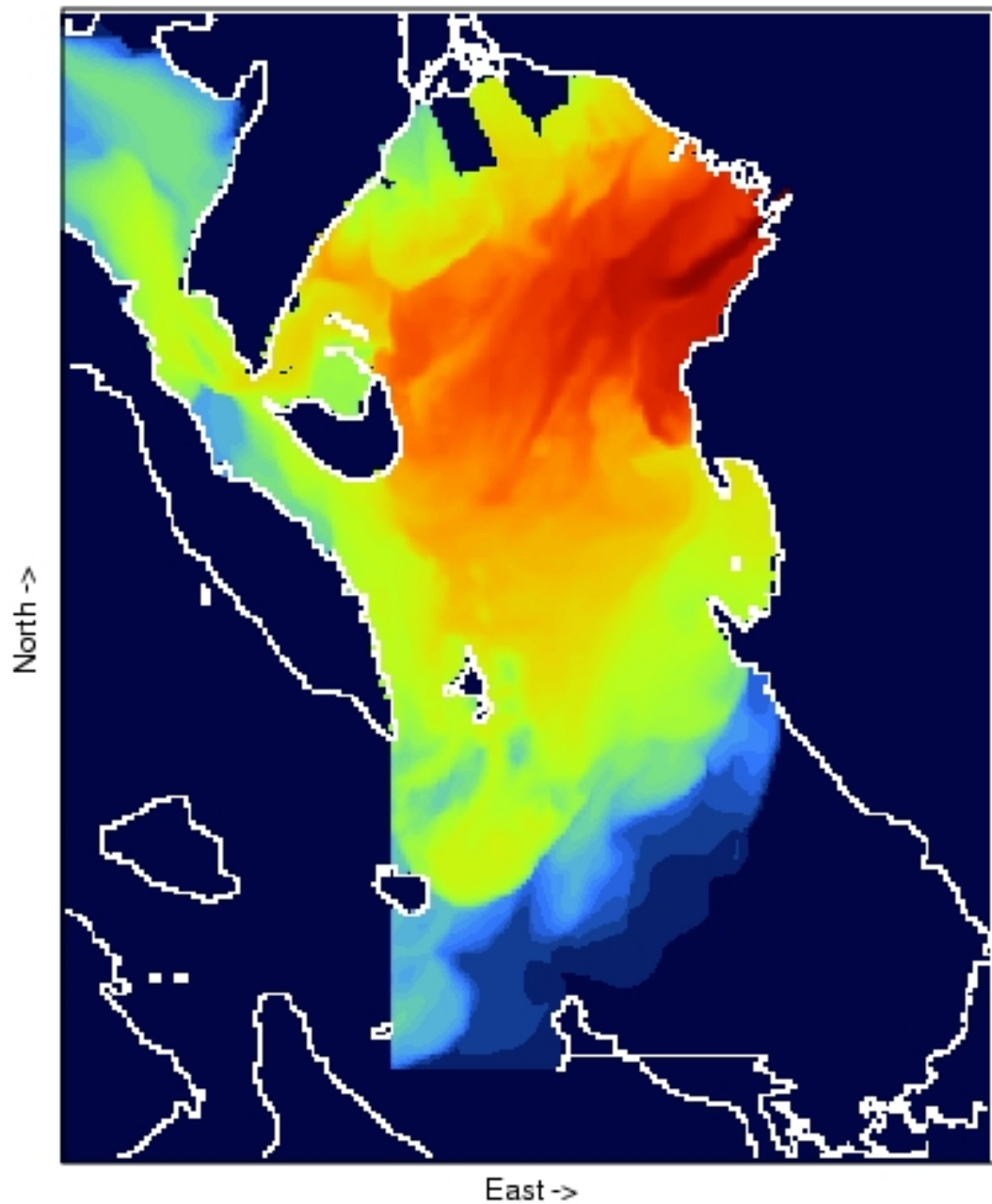


Figure 4: One frame from an animation of color-coded contours of a surface tracer plume.

References

- [Blitz++] <http://www.oonumerics.org/blitz/>
- [Blumberg & Mellor, 1987] Blumberg A.F. and Mellor G.L., 1987. A description of a three-dimensional coastal ocean circulation model. In : N.S. Heaps (Editor), *Three-dimensional Coastal Ocean Models. Coastal and Estuarine Sciences*, Vol. 4, American Geophysical Union, Washington D.C., pp. 1-16.
- [GMT] <http://gmt.soest.hawaii.edu/>
- [Gnuplot] <http://www.gnuplot.info/>
- [GS] <http://www.ghostscript.com/>
- [Huthnance, 1997] Huthnance J.M. (Editor), 1997. Processes in Regions of Freshwater Influence - PROFILE, AS2-CT93-0054. Final Report. POL Internal Document No. 102, Proudman Oceanographic Laboratory, 95 pp.
- [Kdevelop] <http://www.kdevelop.org/>
- [Luyten, 1999] Luyten P.J. (Editor), 1999. COHERNS - Dissemination and exploitation of a coupled hydrodynamical-ecological model for regional and shelf seas, MAS3-CT97-0088. Final Report. MUMM Internal Report, Management Unit of the Mathematical Models, 76 pp.
- [Ncview] http://meteora.ucsd.edu/~pierce/ncview_home_page.html
- [NEdit] <http://www.nedit.org>
- [NetCDF] <http://my.unidata.ucar.edu/content/software/netcdf/index.html>
- [NOS] <http://www.ngdc.noaa.gov/mgg/bathymetry/hydro.html>
- [Proctor, 1997] Proctor R. (Editor), 1997. NOMADS - North Sea Model Advection-Dispersion study, MAS2-CT94-0105. Technical Report IV: Model Intercomparison. POL Internal Document No. 107, Proudman Oceanographic Laboratory, 143 pp.
- [Vis5d+] <http://vis5d.sourceforge.net/>
- [XTide] <http://www.flaterco.com/xtide/>